

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 16, 2003, 15:50:54 : Search time 2492.14 Seconds
(without alignments)
16465.781 Million cell updates/sec

Title: US-09-497-967-44
Perfect score: 1410
Sequence: 1 atgaataataatttagt.....cttattattattatgatga 1410

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rod.*
36: em.htg_mam.*
37: em.htg_vrt.*
38: em.sy.*
39: em.htgo_hum.*
40: em.htgo_mus.*
41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1410	100.0	3026	3	AF324424
2	259.8	18.4	1520	3	AF405431
3	258	18.3	2486	3	AF140273
4	214.6	15.2	1249	3	ICYMANT
C 5	96	6.8	175748	2	AC120669
C 6	96	6.8	180903	2	AC125859
C 7	95.4	6.8	186935	2	AC022322
C 8	92.6	6.6	179553	2	AC024253
C 9	90.8	6.4	131274	2	AC098669
C 10	90.8	6.4	176822	2	AC098669
C 11	85.6	6.1	261604	2	AC099174
C 12	85.4	6.1	127354	2	AC119819
C 13	85.4	6.1	155019	2	AC117014
C 14	81	5.7	35793	5	AY016024
C 15	81	5.7	84472	2	AC098684
C 16	80.2	5.7	132449	9	AL365272
C 17	77.2	5.5	22398	5	FRU271723
C 18	77	5.5	85786	9	AL162582
C 19	76.6	5.4	10115	5	AF397467
C 20	75.8	5.4	183413	2	AC131200
C 21	75.6	5.4	170985	2	AC096032
C 22	74.6	5.3	204259	2	AC110817
C 23	74	5.2	172307	2	AC044842
C 24	72.4	5.1	189461	2	AC119627
C 25	72.4	5.1	191841	2	AC121374
C 26	72	5.1	134558	2	AC125757
C 27	72	5.1	180668	2	AC020857
C 28	69.8	5.0	666	8	AF413050S2
C 29	69.2	4.9	172853	9	AC084361
C 30	68.2	4.8	71553	2	AC096997
C 31	68	4.8	154141	2	AC121042
C 32	68	4.8	162269	9	AC000097
C 33	67.8	4.8	184830	2	AC113220
C 34	67.2	4.8	185994	2	AC002042
C 35	67.2	4.8	300994	2	AC128317
C 36	66.2	4.7	131346	2	AC119558
C 37	66	4.7	138615	2	AC117835
C 38	65.8	4.7	101534	2	AC108576
C 39	65.8	4.7	298283	3	AE003782
C 40	64.8	4.6	177035	2	AC099443
C 41	64.4	4.6	186558	2	AC079031
C 42	64.4	4.6	325808	2	AC121787
C 43	64.2	4.6	187617	2	AC118993
C 44	64	4.5	100726	2	AC116961
C 45	63.8	4.5	177435	2	AC128321

ALIGNMENTS

RESULT 1
AF324424
LOCUS
DEFINITION
Ichthyophthirius multifiliis 3026 bp DNA linear INV 27-FEB-2002
(IAGS2A) gene, complete cds.
ACCESSION
AF324424
VERSION
AF324424.1 GI:12698726
KEYWORDS
SOURCE
ORGANISM
Ichthyophthirius multifiliis.
Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;
Hymenostomatida; Ophryoglenina; Ichthyophthirius.
REFERENCE
1 (bases 1 to 3026)
Lin.Y., Lin.T.L., Wang.C.C., Wang.X., Stieger.K., Klopfleisch.R.
and Clark.T.G.

TITLE Variation in primary sequence and tandem repeat copy number among
i-antigens of *Ichthyophthirius multifiliis*
JOURNAL Mol. Biochem. Parasitol. 120 (1), 93-106 (2002)
MEDLINE 21839613
PUBMED 11849709
REFERENCE 2 (bases 1 to 3026)
AUTHORS Lin, Y., Lin, T.-L. and Clark, T. G.
TITLE Direct Submission
JOURNAL Submitted (27-NOV-2000) Microbiology and Immunology, NYSCVM Cornell
University, Ithaca, NY 14853, USA
FEATURES
source
Location/Qualifiers
1..3026
/organism="Ichthyophthirius multifiliis"
/macronuclear
/serotype="D"
/isolate="G5"
/db_xref="taxon:5932"
1285..2817
/gene="TAG52A"
1285..2817
/gene="TAG52A"
/product="immobilization antigen isoform"
1330..2736
/gene="TAG52A"
/note="52 kDa protein"
/codon_start=1
/transl_table=6
/product="immobilization antigen isoform"
/protein_id="AAK01661.1"
/db_xref="GI:12698727"

gene
mRNA
CDS

BASE COUNT 1177 a 316 c 355 g 1178 t
ORIGIN
Query Match 100.0%; Score 1410; DB 3; Length 3026;
Best Local Similarity 100.0%; Pred. No. 4.1e-237;
Matches 1410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAAAATAATTTAGTAAATGATTTATTTCAATTTATCAATTAATCT 60
Db 1330 ATGAAAAATAATTTAGTAAATGATTTATTTCAATTTATCAATTAATCT 1389
QY 61 GCTAATTCCTGTTGGAACCTGAACCTAACACACGCGGATAAGTTGATCTAGGAAC 120
Db 1390 GCTAATTCCTGTTGGAACCTGAACCTAACACACGCGGATAAGTTGATCTAGGAAC 1449
QY 121 CTGCAAAATGTTGTTAAATGTTAGAAAACCTTTTATTATAAATGCTGCTTTCGTT 180
Db 1450 CTGCAAAATGTTGTTAAATGTTAGAAAACCTTTTATTATAAATGCTGCTTTCGTT 1509
QY 181 CTGGTGCTAGTACGTGACCTTGCCATAAAAAAAGATGCTGCTTAACCAAT 240
Db 1510 CTGGTGCTAGTACGTGACCTTGCCATAAAAAAAGATGCTGCTTAACCAAT 1569
QY 241 CCACCTGCTACTGCTAAATTTAGTACATTAATGTAACGTTAAATGCCCTGCTGTAACGCA 300
Db 1570 CCACCTGCTACTGCTAAATTTAGTACATTAATGTAACGTTAAATGCCCTGCTGTAACGCA 1629
QY 301 ATTGCAGGTGGAGCAACAGATTATCCAGCAATATATCAGCAATGTGTTAATTTAGAAAT 360
Db 1630 ATTGCAGGTGGAGCAACAGATTATCAGCAATATATCAGCAATGTGTTAATTTAGAAAT 1689
QY 361 AATTTTATAAAGAAATGCTCAATTTAATGAGGTGCTAGTACATGACAGCTTGT 420
Db 1690 AATTTTATAAAGAAATGCTCAATTTAATGAGGTGCTAGTACATGACAGCTTGT 1749

QY 421 CCGGTAACACAGATGGTGGTGCAATTTGACTGCTGTAATCCGCTACCATAGTCGCATAA 480
Db 1750 CCGGTAACACAGATGGTGGTGCAATTTGACTGCTGTAATCCGCTACCATAGTCGCATAA 1809
QY 481 TGTAAAGCTGCGATGCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 1810 TGTAAAGCTGCGATGCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1869
QY 541 AGATCAATTCACAGATGTTAAATGTAAGTAAATGTAAGTAAATGTAAGTAAATGTAAGT 600
Db 1870 AGATCAATTCACAGATGTTAAATGTAAGTAAATGTAAGTAAATGTAAGTAAATGTAAGT 1929
QY 601 AATACTCCTTTCAATCCAGGTAAAGTAAATGTAAGTAAATGTAAGTAAATGTAAGTAAAT 660
Db 1930 AATACTCCTTTCAATCCAGGTAAAGTAAATGTAAGTAAATGTAAGTAAATGTAAGTAAAT 1989
QY 661 AATGTTGCTTAAAGCTACTTTAGGTAAATGTAAGTAAATGTAAGTAAATGTAAGTAAAT 720
Db 1990 AATGTTGCTTAAAGCTACTTTAGGTAAATGTAAGTAAATGTAAGTAAATGTAAGTAAAT 2049
QY 721 TGGCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 2050 TGGCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2109
QY 781 TGTACTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 2110 TGTACTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2169
QY 841 AGTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 2170 AGTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2229
QY 901 GCGCTACTTTAGCCAAATTAATGTAATTTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 2230 GCGCTACTTTAGCCAAATTAATGTAATTTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2289
QY 961 GGAGCAACTAATTAATGTAATTTAATGTAATTTAATGTAATTTAATGTAATTTAATGTA 1020
Db 2290 GGAGCAACTAATTAATGTAATTTAATGTAATTTAATGTAATTTAATGTAATTTAATGTA 2349
QY 1021 TTTGATGTAATTAATTTCTAGGAGGAGTGTAGTGTGCAAGCAATGCTCCAGCAATAA 1080
Db 2350 TTTGATGTAATTAATTTCTAGGAGGAGTGTAGTGTGCAAGCAATGCTCCAGCAATAA 2409
QY 1081 GTTAAAGCGCTGTAGCAACTGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Db 2410 GTTAAAGCGCTGTAGCAACTGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2469
QY 1141 GAATGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
Db 2470 GAATGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
QY 1201 TCTGAATGTTAAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
Db 2530 TCTGAATGTTAAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2589
QY 1261 GGTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Db 2590 GGTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2649
QY 1321 CCTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
Db 2650 CCTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2709
QY 1381 TTATTGATTTCTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 1410
Db 2710 TTATTGATTTCTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 2739

RESULT 2
AF405431
LOCUS AF405431 1520 bp mRNA linear INV 27-FEB-2002

DEFINITION Ichthyophthirius multifiliis 52kDa immobilization antigen variant B
protein mRNA, complete cds.
ACCESSION AF405431
VERSION AF405431.1 GI:15290741
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS
BASE COUNT 493 a 255 c 256 g 516 t
ORIGIN
Query Match 18.4%; Score 259.8; DB 3; Length 1520;
Best Local Similarity 55.2%; Pred. No. 6.5e-36;
Matches 800; Conservative 0; Mismatches 547; Indels 102; Gaps 11;
QY 1 ATCAAAATTAATTTTACTAATATTGATTATTTTCATTTATTTATCAATTAATTAATCT 60
DB 41 ATCAAAATTAATTTTAAATATTGATTATTTCTTATTTATTAATGAATTAAGACT 100
QY 61 GCTAATGTCCTGTTGGAATGAACTACACAGCGCGGATAAGTTGATGATCTAGGAAT 120
DB 101 GTTAATGTCCTAATGCTGCAATGCGAATGGATGATATCTGATACAGGACTGCAGAT 160
QY 121 CCTGCAAAATGTTAATTTAGAAAACTTTTATATAATAATGCTGCTGCT----- 174
DB 161 ATAAATCTGTACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 220
QY 175 -----TTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 213
DB 221 CAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 280
QY 214 AAAAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 273
DB 281 CACAAA---GCCGATCTCAACACAGATAAGGTGCTGATGCTAATTTAGCCGCATAATGT 337
QY 274 AACGTTAAATGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 333

DB 338 AGCAACTTATGCTGCTGGCAGTGAAGATGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 397
QY 334 ATCAAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 393
DB 398 CTACATAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 457
QY 394 GCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 453
DB 458 CAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 517
QY 454 GCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 513
DB 518 AATAATCC-----TGAAGTTCCTAATGTTTCTAGCCCTAAATGGTAAATGCTGCTGCT 565
QY 514 GATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 573
DB 566 GCTTGTGCTAAGTAAACA---AGTCTGATCTCAATTAAGACAGGCTGCTGCTGCTGCTGCT 622
QY 574 AACTTTTACTATAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 633
DB 623 GCCACATAATGCTAAGTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 682
QY 634 ACACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 693
DB 683 ATTTATCTTAACTCAATCTCATAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 742
QY 694 ACAATAACCGCATAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 753
DB 743 -----GCAATCCTTCAGCTCAGATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 775
QY 754 ATAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 813
DB 776 TAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 835
QY 814 AATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 873
DB 836 ATGGTTC-----TGGCCCTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 886
QY 874 TATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 933
DB 887 AA---TTCTAATCCAGATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 943
QY 934 TGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 990
DB 944 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1003
QY 991 GAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1050
DB 1004 TAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1060
QY 1051 AGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1110
DB 1061 AGTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1117
QY 1111 ACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1170
DB 1118 AGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1177
QY 1171 GGAACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1230
DB 1178 GGTACATCACTAATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1237
QY 1231 TATACTACAAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1290
DB 1238 TATGCTACAAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1297
QY 1291 AATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1350
DB 1298 AATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1357
QY 1351 G-----ATTTCGCTAATTTTATCAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1401

Db 1358 GCCAGTTCACCTTCCGAAAAATTTTATCAATGTCCTTAATAATTATTTCTTCTATTG 1417

QY 1402 TTATGATGA 1410

Db 1418 TTGTGATGA 1426

RESULT 3

AF140273 2486 bp DNA linear INV 19-MAY-1999

LOCUS Ichthyophthirius multifiliis immobilization antigen precursor

DEFINITION (IAG48) gene, complete cds.

ACCESSION AF140273

SOURCE AF140273.1 GI:4868370

VERSION

KEYWORDS

ORGANISM Ichthyophthirius multifiliis.

Ichthyophthirius multifiliis

Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;

Hymenostomatida; Ophryoglenina; Ichthyophthirius.

1 (bases 1 to 2486)

Clark,T.G., Lin,T.L., Jackwood,D.A., Sherrill,J., Lin,Y. and

Dickerson,H.W.

The gene for an abundant parasite coat protein predicts tandemly

repetitive metal binding domains

Gene 229 (1-2), 91-100 (1999)

99196987

10095108

2 (bases 1 to 2486)

Gaertig,J., Gao,Y., Tishgarten,T., Clark,T.G. and Dickerson,H.W.

Surface display of a parasite antigen in the ciliate Tetrahymena

thermophila

Nat. Biotechnol. (1999) In press

3 (bases 1 to 2486)

Clark,T.G., Lin,T.-L., Jackwood,D.A. and Dickerson,H.W.

Direct Submission

Submitted (29-MAR-1999) Microbiology & Immunology, Cornell

University, College of Veterinary Medicine, Ithaca, NY 14853, USA

Location/Qualifiers

1. .2486

/organism="Ichthyophthirius multifiliis"

/serotype="A"

/isolate="GI"

/db_xref="taxon:5932"

399. .1845

/gene="IAG48"

399. .1845

/gene="IAG48"

/product="immobilization antigen precursor"

433. .1761

/gene="IAG48"

/note="surface protein; 48 kDa i-antigen"

/codon_start=1

/transl_table=6

/product="immobilization antigen precursor"

/protein_id="AAD31283.1"

/db_xref="GI:4868371"

translation="MKYNILLILILSLFNLRAVPCPDGTQAGLTDVGAADLGTG

VNCRPFYNGGAOGEANGNPFANNAAGICVPCQINRVGSVNTNAGDLATATQC

STOCPGTALDDGTVDFDRSRAOCVKCPNPNYNGSGPOGEAPGVQVPAAGAAAGV

AAVTSQCVQLNKNDSPATGAQAANLTCQSNQCPGTGVLDGVLVFNTPSATLGVK

CRNFYNGSGPOGEAPGVQVFAAGAAAGVAAVTSQCVPCQINKNDSPATGAQAANL

ATQCSQCPGTATIQDGVTLVFNSTQCSQCIANYFENGFEAGSKCPVSKTT

PAHAPGNATQATQCLTTCPTAGTVLDDGTSTNFVASATECTKCSAGFFASKTTGTAG

TDNCTCKTLTSGATAKVYAEATQKVOCASTTFKFLSLISLIFSLL"

433. .492

/gene="IAG48"

493. .1758

/gene="IAG48"

/product="immobilization antigen"

896 a 310 c 321 g 959 t

sig_peptide

mat_peptide

BASE COUNT 896 a 310 c 321 g 959 t

ORIGIN

Query Match 18.3%; Score 258; DB 3; Length 2486;

Best Local Similarity 56.8%; Pred. No. 1.2e-35;

Matches 666; Conservative 0; Mismatches 395; Indels 111; Gaps 6;

QY 344 GTGTTAATTTGAGAAATTAATTTTATATAATGAAAATGTCCTCAAAATTTTAAATGCGAGTGCTA 403

Db 599 GTGCTGCTTAAGGAGAAGCTAATGGTAATTAACCTTTCCGCGACAAATAATGCTGCTAGAG 658

QY 404 CTACATGCACAGCTTTCGCGGTAAACAGAGTTGGTGGTGCATTTGACTGCTGGTAAATGCCG 463

Db 659 GTATATGTTGTTACCATGCCAAATAAACAGAGTAGGCTCTGTATACCAATGCAGTGACTTAG 718

QY 464 CTACATAGTCGCATAATGTAACGTCGCATGCTCTGCTGCTAGTGCACCTGATGATGAG 523

Db 719 CTACTTTAGCCACATAATGTCAGTACTTAATGCTCTGTCGACCTGTCATGATGAG 778

QY 524 TAACCTACTGATTATGTTAGATCAATTCACAGAAATGTTAAATGTAGACTTAACCTTTTACT 583

Db 779 TGACAGATGTTTGTGATAGATCAGCCGCATAATGTTTAATGCAAACTTAACCTTTTACT 838

QY 584 ATAATGGTAAATAGGTAAATACCTCTTTCATTCAGGTAAAGTTAATGCACACCTTGTG 643

Db 839 ATAATGGTGGTTCTCCTTAAGGTGAAGCTCCTGGCGTTTAAAGTTTTTGTCTGCTGCTGCTG 898

QY 644 CGGCAATTTAAACCTGCT-----AATG 664

Db 899 CGGCTGCAGGTGTTGCTGCCGTACTAGTTAATGTGACCTTGCACACTAAACAAACAG 958

QY 665 TTGCTTAAGCTACTTTAGTGAATGATGCTACAATAAACGCGCATATGTAACTTGCATGCC 724

Db 959 ATTCTCTGCCACTCGCAGGTGCCCTAAGCTAATTAGCCACATAATGTAGCAATTAATGTC 1018

QY 725 CTGATGCTACTATAAGTCTGCTGCTGAGT---AAATAATTTGGTAGGACCAAACTGAT 781

Db 1019 CTACTGGCAGCTGCTACTTGTATGATGATGAGTGCACCTTGTTTTAAATACATCAGCCACAT 1078

QY 782 GTACTAATTTGCTCCTTAACCTTTTACAATAATAATGCTCTAAAT----- 826

Db 1079 GTGTTAAATGACAGCTCACTTTTACTATAATGTTGGTGTCTCCTTAAGTGAAGTCTCTG 1138

QY 827 -----TCAATCCAGGTAAATAGTACAT 847

Db 1139 GCGTTTAAGTTTTCGTGCTGCCGTGCAGGTGTTGTCGCCGTACTAGTTAAT 1198

QY 848 GCCTACCTTGGCCAGCAATAAAGATTATGGTGTGAAGCCACTGCAGGTGGTGCCGCTA 907

Db 1199 GTGTACCTTGGCCAAATAAACAAAAAGATCTCTCT---GCCACTGCAGGTGCCCTAAGTGA 1255

QY 908 CTTTAGCCAAATAATGTAATATTGCATGCCCTGAGTGGTACTGCAATTTGCTAGTGGAGCAA 967

Db 1256 ATTTAGCCACATAATGTCAGTACTTAATGTCCAACTGGCAGTGCATTTCAAGACGAGTGA 1315

QY 968 CTAATTTATGTAATATATAACAGAAATGCTAAATTTGCTGCTTAACCTTTTATTTTATG 1027

Db 1316 CACTTGTTTTATGTAATTCATCCACATAATGTTCTTAATGCAATTTGCTAATTTTATTTA 1375

QY 1028 GTAATAATTTCTAGGAGGAAGTAGTAGTGAAGCATGTCCGCAATAAAGTTTAAAG 1087

Db 1376 ATGGTAATTTTCAAGCAGGTAAGAGTTAATGTTTAAAGTGTCCAGTAGTAAGTAACT---A 1432

QY 1088 GCGCTGTAGCAACTGCAGGTGGTACTGCTACTTAAATTTAATGCAATAATGTCCTTTGAATGCC 1147

Db 1433 CTCACGACATGCTCCAGGTAATGCTACTTAAGCCACATAATGTTTGACACATGTC 1492

QY 1148 CTGCTGGTACTGTACTACCGTAGGAACAACATCTACTTATTAATAAGCAGCATCTGAAT 1207

Db 1493 CTGCTGGTACAGTACTTGTATGATGGAACATCACTAATTTTGTAGCTTCGCAACTGAAT 1552

QY 1208 GTGTTAAATGTGCTGCCAACTTTTATACATACAAATAAATGATTTGGGTAGCAGGTATTG 1267

Db 1553 GTACTAATTTGTTCTGCTGGCTTTTTCATCAAAAACAACACTGGTTTACAGCAGGACTG 1612

QY 1268 ATACATGTACTAGTTGTAATAAAAAATTAACCTTCTGGCGCTGAAGCTAATTTACCTGAAT 1327

Db 1613 ATACATGACTGAATGTACTAAAAAATAAATCTCTGGTGCCACAGCTAAAGTATATGCTG 1672

Qy 1328 CTGCTAAAAAATATATAATG-----TGATTTTCGCTAATTTTATCAATTTTCCT 1378

Db 1673 AAGTACTCAAAAGTATAATGCCCTCCACTACTTTTCGCTAAATTTTATCGATTTCCT 1732

Qy 1379 TATTATTGATTCTTATTATTATTATATGATGA 1410

Db 1733 TATTATTATTCTTCTATTATTATGATGA 1764

RESULT 4

ICYMANT

LOCUS

DEFINITION Ichthyophthirius multifiliis 1249 bp mRNA linear INV 18-SEP-1998

ACCESSION M92907

VERSION M92907.1 GI:3628568

KEYWORDS

SOURCE

ORGANISM

Ichthyophthirius multifiliis.

Ichthyophthirius multifiliis

Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;

Hymenostomida; Ophryoglenina; Ichthyophthirius.

REFERENCE

AUTHORS

TITLE

Lin, T.L. and Dickerson, H.W.

Purification and partial characterization of immobilization

antigens from Ichthyophthirius multifiliis

J. Protozool. 39 (4), 457-463 (1992)

JOURNAL

MEDLINE

PUBMED

93020590

1383510

REFERENCE

AUTHORS

TITLE

Clark, T.G., McGraw, R.A. and Dickerson, H.W.

Developmental expression of surface antigen genes in the parasitic

ciliate Ichthyophthirius multifiliis

Proc. Natl. Acad. Sci. U.S.A. 89 (14), 6363-6367 (1992)

JOURNAL

MEDLINE

PUBMED

92335298

1631132

REFERENCE

AUTHORS

TITLE

Clark, T.

Direct Submission

Submitted (06-JUN-1992) Microbiology and Immunology, Cornell

University, Ithaca, NY 14853, USA

JOURNAL

REFERENCE

AUTHORS

TITLE

Clark, T.

Direct Submission

Submitted (18-SEP-1998) Microbiology and Immunology, Cornell

University, Ithaca, NY 14853, USA

REMARK

Sequence update by submitter

On Sep 18, 1998 this sequence version replaced gi:159289.

FEATURES

source

1. .1249

/organism="Ichthyophthirius multifiliis"

/strain="Georgia"

/db_xref="taxon:5932"

/dev_stage="trophont"

<1. .1190

/note="48 kDa"

/citation=[2]

/citation=[11]

/codon_start=3

/evidence=experimental

/transl_table=6

/product="Immobilization antigen precursor"

/db_xref="GI:3628568"

/translation="AVPCPDGTQTQAGLTDVGADLGTVCNCRPNFYNGGAAQGBAN

RSQPFPAANRAGICVPCQINRVGSVYNAGDLATLQCTSTOCPTGTLDDGVDVDF

NSAQCCKCPNRYNGSGPQGPVQVFAAGAAAAGAAVTSQCPQLNKNDSPA

TAGAQLANTQCSNOCPGTGLVDGTVLVFNTSATLCKVCRPNFYNGSGSPQGEAPGV

QVFAAGAAAGAAVTSQCPVPCQINKNDSPATAGAAANLATQCTSTOCPTGTLQDQVT

LVSFNSQCSQCIANYFFNGNLEAKRSCLKCPVSKTTPAHAPGNATQATQCLITC

PAGTVLDDGTSTNFVASATECTKCSAGFFASKTTGTAGTDTCTECTKLTSGATAVN

IHQ"

sig_peptide

mat_peptide

repeat_region

BASE COUNT 348 a 243 c 257 g 401 t

ORIGIN

Query Match 15.2%; Score 214.6; DB 3; Length 1249;

Best Local Similarity 55.4%; Pred. No. 5.4e-28;

Matches 591; Conservative 0; Mismatches 374; Indels 102; Gaps 5;

Qy 344 GTGTTAAATGTAGAAATTAATTTTATTAATGAAATGCTCCAAATTTTAAATGAGGTGCTA 403

Db 112 GTGCTGCTTAAAGGAAGCTAATGTAATTAACCTTTTCGACGCAANAATGCTGCTAGAG 171

Qy 404 GTACATGCACAGCTTGTCCGGTAAACAGAGTTGGTGGTGCATTTGACTGTGTAATGCCG 463

Db 172 GTATATGTGTACCATGCCAANAATAACAGAGTAGGCTCTGTTACCAATGCAGGTGACTTAG 231

Qy 464 CTACCATAGTCGCATAATGTAAAGTCGCATGCTCTACTGCTACTGTCACATTTGATGGAG 523

Db 232 CTACTTTAGCCACATAATGCAGTACTTAATGTCTTACTGGCACTGCACCTTGATGGAG 291

Qy 524 TAACTACTGATATGTAGATCATTCACAGAAATGTGTTAAATGTAGACTTAACCTTTTACT 583

Db 292 TGACAGATGTTTTTGATAGATCAGCCGCAATAATGTGTTAAATGCAACCTAACTTTTACT 351

Qy 584 ATAAATGGTAAATATGTAATACTCTCTTCAATCCAGGTAAAAGTTAATGCACACACTTTGTC 643

Db 352 ATAAATGGTGGTCTCTCTTAAGGTGAAGCTCCTGGCGCTTAAGTTTGTGCTGGTGGCTG 411

Qy 644 CGCAATTAACACCTGCT-----ATG 664

Db 412 CGCTGCAGGTGTGCTGCCGTGTACTAGTTAATGTGTACTTGTCCCACTAAACAAAAAG 471

Qy 665 TTGCTTAAAGTACTTTTAGTGTAAATGATGCTACAAATACCGCATANTGTAAGTTGATGCC 724

Db 472 ATTCTCCTGCCACTGCAGCTGCGCTAAGCTAAATTTAGCCACATAATGTAGCAATTAATGTC 531

Qy 725 CTGATGGTACTATAAGTGTGCTGCTGGAGT---AAATAATTTGGGTAGCACAAAAACACTGAAT 781

Db 532 CTACTGGCACTGTACTTTGATGATGGAGTGACACTTGTCTTTTAAATCATCAGCCACATTAAT 591

Qy 782 GTACTAATTTGTGCTCCTAACTTTTACAATAATAATGCTCCTAATTT-----826

Db 592 GTGTTAAATGCAGACCTAACTTTTACTATAATGTTGTTCTGCTTAAAGGTGAAGCTCCTG 651

Qy 827 -----TCAATCCAGCTAATAGTACAT 847

Db 652 GCCTTAAAGTTTTGTGCTGGTGGTGGCGCTGCAGGTGTGCTGCCGTGTACTAGTTAAT 711

Qy 848 GCCTACTTTGCCAGCAAAATAAGATTATGGTGTCAAGCCACTGCAGGTGGTGGCGGCTA 907

Db 712 GTGTACTTTGCCAANAATAACAAAAACGATTTCTCCT---GCCACTGCAGGTGCCTAGCTA 768

Qy 908 CTTTAGCCAAAATAATCTAATATTGTGCATGCCCTGATGGTACTGCAATTTGTTAGTGGAGCAA 967

Db 769 ATTTAGCCACATAATGCAGTACTTAATGTCCAACCTGGCAATTCGAATTAAGACGGAGTGA 828

Qy 968 CTAATTTATGTAATTTATTAACAGAAATGCTAANAATGTGCTGCTACTACTTTTATTTTATG 1027

Db 829 CACTGTTTTTAGTAATTCATCCACATAATGTTCTTAATGTCATGCTGCTAATTTTACTTTT 888

Qy 1028 GTAATAATTTCTAGGCAAGAGTAGTAGATGCAAGCATGTGCAGCAAAATAAGTTTAAAG 1087

Db 889 ATGCTAATCTCGAAGCAGGTAAAGTTAATGTTTAAAGTGTCCAGTAAAGTAAACT---A 945

Qy 1088 GCCTGTAGCAACTGCAGGTGGTGTGCTACTTTTAATTCATTAATGTCCTTGAATGCC 1147

Db 946 CTCACGACATGCTCCAGGTATATACGTCTACTTAGCCGCACATAATGTTTGACCACATGTC 1005
QY 1148 CTGCTGGTACTGTACTACCCGATGGAACAACATCTACTTATATAAATGAAGCAGCATCTGAAT 1207
Db 1006 CTGCTGGTACAGTACTGTGATGGACATCAACTAAATTTGTAGCTCCGCCAACTGAAT 1065
QY 1208 GTGTTAAATGTGTCGCCAACTTTTATCTACTACAAAATAAATGATTTGGGTAGCAGGTATTG 1267
Db 1066 GTACTAAATGTTCTGCTGGCTTTTGTGCATCAAAAACAACACTGTTTTCACAGCAGGTACTG 1125
QY 1268 ATACATGTACTAGTCTATATAAAAAATTAACCTTCTGGCGCTGAAGCT 1314
Db 1126 ATACATGTACTGAATGTACTAAAAAATTAACCTTCTGGTGCACAGCT 1172

RESULT 5

AC120669/c

LOCUS AC120669 175748 bp DNA linear HTG 23-JUL-2002
DEFINITION Rattus norvegicus clone CH230-22G10, *** SEQUENCING IN PROGRESS
***, 46 unordered pieces.

AC120669

AC120669.2 GI:21902861

HTG: HTGS_PHASE1.

Rattus norvegicus.

Rattus norvegicus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

AUTHORS

1 (bases 1 to 175748)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
Alsbrooks, S.L., Amarantunge, H.C., Are, J.R., Ayelle, M., Banks, T.,
Barbacia, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,
Bouck, S., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabrelli, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,
Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,
Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mathiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
Miner, G., Mear, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, J., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenwo, S., Ogih, M., Okwuonu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shooshitari, N., Sisson, I.,
Sutton, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wlarczyk, R., Wooden, S., Worley, K.,
Weinstock, G. and Gibbs, R.

TITLE

JOURNAL

AUTHORS

2 (bases 1 to 175748)

Unpublished

Worley, K.C.

TITLE
JOURNALREFERENCE
AUTHORSTITLE
JOURNAL

COMMENT

Direct Submission
Submitted (09-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 175748)
Worley, K.C.
Direct Submission
Submitted (23-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 18, 2002 this sequence version replaced gi:20514530.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GXL

Center clone name: CH230-22G10

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 130722 bases at least Q40

Consensus quality: 138187 bases at least Q30

Consensus quality: 143539 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 46 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 1312: contig of 1312 bp in length

* 1313 1412: gap of unknown length

* 1413 2562: contig of 1150 bp in length

* 2563 2662: gap of unknown length

* 2663 3966: contig of 1304 bp in length

* 3967 4066: gap of unknown length

* 4067 5262: contig of 1196 bp in length

* 5263 5362: gap of unknown length

* 5363 6793: contig of 1431 bp in length

* 6794 6893: gap of unknown length

* 6894 8272: contig of 1379 bp in length

* 8273 8372: gap of unknown length

* 8373 9620: contig of 1248 bp in length

* 9621 9720: gap of unknown length

* 9721 11136: contig of 1416 bp in length

* 11137 11236: gap of unknown length

* 11237 12579: contig of 1343 bp in length

* 12580 12679: gap of unknown length

* 12680 14466: contig of 1787 bp in length

* 14467 14566: gap of unknown length

* 14567 16613: contig of 2047 bp in length

* 16614 16713: gap of unknown length

* 16714 18834: contig of 2121 bp in length

* 18835 18934: gap of unknown length

* 18935 20996: contig of 2062 bp in length

* 20997 21096: gap of unknown length

* 21097 22643: contig of 1547 bp in length

* 22644 22743: gap of unknown length

* 22744 25384: contig of 2641 bp in length

* 25385 25484: gap of unknown length

* 25485 27048: contig of 1564 bp in length

* 27049 27148: gap of unknown length

* 27149 29221: contig of 2073 bp in length

* 29222 29321: gap of unknown length

* 29322 30973: contig of 1652 bp in length

* 30974 31073: gap of unknown length

* *

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE AUTHORS

1 (bases 1 to 180903)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L., Amarantunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbarella, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., BuHay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, K.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, M.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Hernandez, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulsegad, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wlezyk, R., Woodson, S., Worley, K., Wu, C., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

TITLE JOURNAL

Direct Submission
Unpublished
2 (bases 1 to 180903)
Worley, K.C.
Direct Submission
Submitted (02-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE AUTHORS TITLE JOURNAL

3 (bases 1 to 180903)
Worley, K.C.
Direct Submission
Submitted (23-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GWUQ
Center clone name: CH230-35H21
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 117810 bases at least Q40
Consensus quality: 124839 bases at least Q30
Consensus quality: 130161 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 69 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1018: contig of 1018 bp in length
* 1019 1118: gap of unknown length
* 1119 2397: contig of 1279 bp in length
* 2398 2497: gap of unknown length
* 2498 3536: contig of 1039 bp in length
* 3537 3636: gap of unknown length
* 3637 5038: contig of 1402 bp in length
* 5039 5138: gap of unknown length
* 5139 6416: contig of 1278 bp in length
* 6417 6516: gap of unknown length
* 6517 7882: contig of 1266 bp in length
* 7883 8056: contig of 1174 bp in length
* 8057 9156: gap of unknown length
* 9157 10223: contig of 1067 bp in length
* 10224 10324: gap of unknown length
* 10325 11849: contig of 1526 bp in length
* 11850 11949: gap of unknown length
* 11950 13459: contig of 1510 bp in length
* 13460 13560: gap of unknown length
* 13561 15061: contig of 1502 bp in length
* 15062 15161: gap of unknown length
* 15162 16401: contig of 1240 bp in length
* 16402 16501: gap of unknown length
* 16502 17502: contig of 1001 bp in length
* 17503 18331: contig of 1229 bp in length
* 18332 18931: gap of unknown length
* 18932 20333: contig of 1202 bp in length
* 20334 20333: gap of unknown length
* 20334 21645: contig of 1412 bp in length
* 21646 21745: gap of unknown length
* 21746 23608: contig of 1863 bp in length
* 23609 23708: gap of unknown length
* 23709 25221: contig of 1513 bp in length
* 25222 25321: gap of unknown length
* 25322 26764: contig of 1443 bp in length
* 26765 26864: gap of unknown length
* 26865 28249: contig of 1385 bp in length
* 28250 28349: gap of unknown length
* 28350 30014: contig of 1665 bp in length
* 30015 30114: gap of unknown length
* 30115 32030: contig of 1916 bp in length
* 32031 32130: gap of unknown length
* 32131 35243: contig of 3113 bp in length
* 35244 35344: gap of unknown length
* 35344 37525: contig of 2182 bp in length
* 37526 37625: gap of unknown length
* 37626 39405: contig of 1780 bp in length
* 39406 39505: gap of unknown length
* 39506 41253: contig of 1748 bp in length
* 41254 41353: gap of unknown length
* 41354 43191: contig of 1838 bp in length
* 43192 43291: gap of unknown length
* 43292 43918: contig of 1627 bp in length
* 43919 45018: gap of unknown length
* 45019 46790: contig of 1772 bp in length
* 46791 46890: gap of unknown length
* 46891 48767: contig of 1876 bp in length
* 48767 48867: gap of unknown length
* 48867 50304: contig of 1438 bp in length
* 50305 50404: gap of unknown length

[illegible]

[illegible]

```
QY 759 TTGGTAGCACAAACACTGAATGTACTAATTTGTCTCCCTAACTTTTACAAATAAATGC 818
Db 7830 TCGTCTGCTC-----TGCTGCTGCTACTGTTATTACTGCGATTCTGTTGCTACTGCTGC 7777
QY 819 TCCTAAATTCANTCAGGTAATAGTACATGCCCTACCTTGCCAGCAAAATAAAGATTATGG 878
Db 7776 TGTTCCTGTTGCTGCTGCTATTTGCTGCTGTTGTTAAATGTGATACATTGCTGTTGCTGA 7717
QY 879 TGCTGAAGCCACTGACAGTGGTCCGCTACTTTAGCCAAATATGTAATATTGCATGCC 938
Db 7716 TGCTGCTGTTGCTATTTGCTGCTGCTACTGTTGCTATGTTGCTGTTGTTGTTGTTGCTG 7657
QY 939 TGATGCTACGCAATGCTAGTGGAGCAACTAATTAATGAATATATAAACAGAAATGCT 998
Db 7656 TCTTTGTTGCTGCTGCTGCTACTGATG---CTGTGGCTGCTATTATTGTTGCTGCTGCT 7600
QY 999 AAATGTTGCTGCTAACTTTTATTTGATGTAATATTTCTAGGAGGAAGTAGATG 1058
Db 7599 TGCTGTTGTTACTGCTGCTATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTAT 7540
QY 1059 CAAGCATGTCAGCAAAATAAGTTTAAAGCGCTGTAGCACTGAGGTGGTACTGCTAC 1118
Db 7539 TGCTGTTGCTGCTGCTATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7480
QY 1119 TTTAATGTCATATGTCCTTGAATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1178
Db 7479 TACTGTTGTTGCTACTGCTATGTAATGATGTTGCTGCTGCTGCTGCTGCTGCTGCTAT 7420
QY 1179 ATCTACTATATAAGCAGCATCTGAATGCTGTTAAAGTCTGCTGCTGCTGCTGCTGCT 1238
Db 7419 TACTACTGCTACTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7360
QY 1239 AAATAAATGATGTTGGTAGCAGGTATTGATACATGCTACTGCTGCTGCTGCTGCTGCT 1298
Db 7359 TGCTGATTTTGTGCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7300
QY 1299 TTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1358
Db 7299 TGTTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7241
QY 1359 TAATTTTATCAATTCCTATTATTTGATTTCTTATTTATTTATTTATGATG 1409
Db 7240 TTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 7190

RESULT 8
AC024253
LOCUS AC024253
DEFINITION Homo sapiens chromosome 6 clone RP11-758C19, WORKING DRAFT
SEQUENCE, 12 unordered pieces.
ACCESSION AC024253.4 GI:9959959
VERSION AC024253.4
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 179553)
JOURNAL Waterston,R.H.
REFERENCE The sequence of Homo sapiens clone
AUTHORS 2 (bases 1 to 179553)
JOURNAL Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (28-FEB-2000) Genome Sequencing Center, Washington
MO 63108, USA
COMMENT On Sep 1, 2000 this sequence version replaced gi:8568958.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
```

```
----- Project Information -----
Center project name: H_NH0758C19
----- Summary Statistics -----
Sequencing vector: M13; 100%
Chemistry: Dye-primer Ex; 100% of reads
Assembly: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 175235 bases at least Q40
Consensus quality: 176522 bases at least Q30
Consensus quality: 177171 bases at least Q20
Insert size: 198000; agarose-fp
Insert size: 178453; sum-of-contigs
Quality coverage: 4.48 in Q20 bases; agarose-fp
Quality coverage: 4.99 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 7554: contig of 7554 bp in length
* 7555 7654: gap of unknown length
* 7655 16349: contig of 8695 bp in length
* 16350 16449: gap of unknown length
* 16450 26475: contig of 10026 bp in length
* 26476 26575: gap of unknown length
* 26576 35880: contig of 9305 bp in length
* 35881 35980: gap of unknown length
* 35981 62010: contig of 26030 bp in length
* 62011 90565: contig of 28455 bp in length
* 90566 117183: gap of unknown length
* 117184 117283: contig of 26518 bp in length
* 117284 163019: gap of unknown length
* 163020 163119: contig of 45736 bp in length
* 163120 166938: contig of 3819 bp in length
* 166939 167038: gap of unknown length
* 167039 169312: contig of 2274 bp in length
* 169313 169412: gap of unknown length
* 169413 171639: contig of 2227 bp in length
* 171640 171739: gap of unknown length
* 171740 179553: contig of 7814 bp in length.
*
* Location/Qualifiers
* 1..179553
* /organism="Homo sapiens"
* /db_xref="taxon:9606"
* /chromosome="6"
* /clone="RP11-758C19"
* 1..7554
* /note="assembly_name:Contig10"
* 7655..16349
* /note="assembly_name:Contig11"
* 16450..26475
* /note="assembly_name:Contig12"
* 26576..35880
* /note="assembly_name:Contig13"
* 35981..62010
* /note="assembly_name:Contig14"
* 62111..90565
* /note="assembly_name:Contig15"
* 90666..117183
* /note="assembly_name:Contig16"
* 117284..163019
* /note="assembly_name:Contig17"
* 163120..166938
* /note="assembly_name:Contig18"
* clone_end:T7
* vector_side:right
```

[illegible]


```

Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenwo, S., Oguh, M., Okwuonu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojebokan, I., Rolfe, M., Ruiz, S., Savary, G.,
Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,
Sodergren, E., Sonaike, F., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tatney, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 176822)
Worley, K.C.
Direct Submission
Submitted (09-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 176822)
Worley, K.C.
Direct Submission
Submitted (12-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 11, 2002 this sequence version replaced gi:18846108.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GGZV
Center clone name: CH230-9811
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 115453 bases at least Q40
Consensus quality: 125501 bases at least Q30
Consensus quality: 133023 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  * consists of 77 contigs. The true order of the pieces
  * is not known and their order in this sequence record is
  * arbitrary. Gaps between the contigs are represented as
  * runs of N, but the exact sizes of the gaps are unknown.
  * This record will be updated with the finished sequence
  * as soon as it is available and the accession number will
  * be preserved.
  *
  * 1053: contig of 1052 bp in length
  * 1152: gap of unknown length
  * 1153: contig of 1027 bp in length
  * 2179: gap of unknown length
  * 2280: contig of 1076 bp in length
  * 3355: gap of unknown length
  * 3456: contig of 1379 bp in length
  * 4834: gap of unknown length
  * 4935: contig of 1288 bp in length
  * 6222: gap of unknown length
  * 6223: contig of 1078 bp in length
  * 6323: gap of unknown length
  * 7401: contig of 1307 bp in length
  * 7500: gap of unknown length
  * 8808: contig of 1359 bp in length
  * 8907: gap of unknown length
  * 10267: contig of 1296 bp in length
  * 10367: gap of unknown length
  * 11662: contig of 1296 bp in length
  * 11663: gap of unknown length
  *
  * 11763: contig of 1050 bp in length
  * 12812: gap of unknown length
  * 12913: contig of 1224 bp in length
  * 14136: gap of unknown length
  * 14137: contig of 1210 bp in length
  * 14237: gap of unknown length
  * 15447: contig of 1320 bp in length
  * 15547: gap of unknown length
  * 16666: contig of 1734 bp in length
  * 16966: gap of unknown length
  * 18700: contig of 1558 bp in length
  * 18800: gap of unknown length
  * 20358: contig of 1567 bp in length
  * 20359: gap of unknown length
  * 20459: contig of 2390 bp in length
  * 22026: gap of unknown length
  * 22126: contig of 1231 bp in length
  * 24516: gap of unknown length
  * 24616: contig of 1230 bp in length
  * 25847: gap of unknown length
  * 25947: contig of 1007 bp in length
  * 27177: gap of unknown length
  * 27277: gap of unknown length
  * 28284: gap of unknown length
  * 28384: contig of 1612 bp in length
  * 29996: gap of unknown length
  * 30096: contig of 1364 bp in length
  * 31460: gap of unknown length
  * 31559: contig of 1365 bp in length
  * 32925: gap of unknown length
  * 33024: contig of 1049 bp in length
  * 34073: gap of unknown length
  * 34173: contig of 1252 bp in length
  * 35425: gap of unknown length
  * 35525: contig of 1455 bp in length
  * 36980: gap of unknown length
  * 37080: contig of 1548 bp in length
  * 38628: gap of unknown length
  * 38728: contig of 1484 bp in length
  * 38729: gap of unknown length
  * 40212: contig of 1740 bp in length
  * 40312: gap of unknown length
  * 42052: contig of 1185 bp in length
  * 42153: gap of unknown length
  * 43337: contig of 2089 bp in length
  * 43438: gap of unknown length
  * 45227: contig of 1263 bp in length
  * 45627: gap of unknown length
  * 46889: contig of 1681 bp in length
  * 46899: gap of unknown length
  * 48670: contig of 1300 bp in length
  * 48671: gap of unknown length
  * 48771: contig of 1219 bp in length
  * 50070: gap of unknown length
  * 50171: contig of 1656 bp in length
  * 51390: gap of unknown length
  * 53145: contig of 1375 bp in length
  * 53246: gap of unknown length
  * 54620: contig of 2200 bp in length
  * 54821: gap of unknown length
  * 54821: contig of 2045 bp in length
  * 56920: gap of unknown length
  * 59065: contig of 1547 bp in length
  * 59165: gap of unknown length
  * 60712: contig of 1064 bp in length
  * 60713: gap of unknown length
  * 61876: contig of 2802 bp in length
  * 61976: gap of unknown length
  * 64778: contig of 1914 bp in length
  * 64779: gap of unknown length
  * 66792: contig of 1683 bp in length
  * 66892: gap of unknown length
  * 66993: contig of 1459 bp in length
  * 68575: gap of unknown length
  * 68576: gap of unknown length
  * 70134: contig of 1988 bp in length
  * 70135: gap of unknown length
  * 72222: contig of 1988 bp in length
  * 70235: gap of unknown length
  *

```



```

Center clone name: CH230-123D20
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; version 0.990329
Assembly program: Phrap; 100% of reads
Consensus quality: 73854 bases at least Q40
Consensus quality: 79579 bases at least Q30
Consensus quality: 83355 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  consists of 57 contigs. The true order of the pieces
  is not known and their order in this sequence record is
  arbitrary. Gaps between the contigs are represented as
  runs of N, but the exact sizes of the gaps are unknown.
  * This record will be updated with the finished sequence
  as soon as it is available and the accession number will
  be preserved.
*
1 1478: contig of 1478 bp in length
* 1479 1578: gap of unknown length
* 1579 2917: contig of 1339 bp in length
* 2918 3017: gap of unknown length
* 3018 4096: contig of 1079 bp in length
* 4097 4196: gap of unknown length
* 4197 5642: contig of 1446 bp in length
* 5643 5742: gap of unknown length
* 5743 6930: contig of 1197 bp in length
* 6940 7039: gap of unknown length
* 7040 8481: contig of 1442 bp in length
* 8482 8581: gap of unknown length
* 8582 10096: contig of 1515 bp in length
* 10097 10196: gap of unknown length
* 10197 11452: contig of 1256 bp in length
* 11453 11552: gap of unknown length
* 11553 12648: contig of 1096 bp in length
* 12649 12748: gap of unknown length
* 12749 14133: contig of 1385 bp in length
* 14134 14233: gap of unknown length
* 14234 15511: contig of 1318 bp in length
* 15552 15651: gap of unknown length
* 15652 17024: contig of 1373 bp in length
* 17025 17124: gap of unknown length
* 17125 18147: contig of 1023 bp in length
* 18148 18247: gap of unknown length
* 18248 19923: contig of 1676 bp in length
* 19924 20023: gap of unknown length
* 20024 21045: contig of 1022 bp in length
* 21046 21145: gap of unknown length
* 21146 22292: contig of 1147 bp in length
* 22293 22392: gap of unknown length
* 22393 23545: contig of 1153 bp in length
* 23546 23645: gap of unknown length
* 23646 25288: contig of 1643 bp in length
* 25289 25388: gap of unknown length
* 25389 26388: contig of 1600 bp in length
* 26389 27088: gap of unknown length
* 27089 29136: contig of 2048 bp in length
* 29137 29236: gap of unknown length
* 29237 30361: contig of 1125 bp in length
* 30362 30461: gap of unknown length
* 30462 31728: contig of 1267 bp in length
* 31729 31828: gap of unknown length
* 31829 33326: contig of 1498 bp in length
* 33327 33426: gap of unknown length
* 33427 34972: contig of 1546 bp in length
* 34973 35072: gap of unknown length
* 35073 37098: contig of 2026 bp in length
* 37099 37198: gap of unknown length
* 37199 38876: contig of 1678 bp in length
* 38877 40520: contig of 1544 bp in length
* 38977 40620: gap of unknown length
* 40521 40521: gap of unknown length

```


